(12) UK Patent Application (19) GB (11) 2 158 626 A

(43) Application published 13 Nov 1985

(21) Application No 8511780

(22) Date of filing 9 May 1985

(30) Priority data

(31) 8412135 (32) 11 May 1984 (33) GB

(71) Applicant
Wang Yong-Min,
Computation Centre, Henan Province, China

(72) Inventor
Wang Yong-Min

(74) Agent and/or Address for Service
Marks & Clerk,
57/60 Lincoln's Inn Fields, London WC2A 3LS

(51) INT CL[4]
H03M 11/00

(52) Domestic classification
G4H 13D KU

(56) Documents cited
GB A 2125197
GB A 2118749
GB A 2066990
GB A 2060231
GB A 2030335
GB 1398882
US 3950734

(58) Field of search
G4H

(54) Encoding chinese and like characters and keyboard therefor

(57) The characters are broken in roots, strokes and patterns. Some basic roots, according to their frequency distributions both in the constitution of characters and in practical usage, are selected optimally and classified in accordance with their inner links and compatible relations, and then arranged on the 25 keys of a standard keyboard in accordance with the frequency or probability of using the keys and the fingering of the keyboard. Keys can be provided corresponding to the to the strokes constituting Chinese characters, a predetemined number of the keys being operated in a sequence corresponding to the sequence of writing the character by hand.

key names	key-code	stroke-root	main roots	derivative roots
王	G(11)	一	干千夫戋主壬	[illegible]
土	F(12)	二	士十丰毛	
大	D(13)	三ア𠂇	厂石戊尢甘且井	犬戶彐非丰手長
木	S(14)		寸西	束
廿	A(15)	七匚丂	丁工开戈	廿廿厶弋
目	H(21)	丨卜⺊	上止卢	广止且
日	J(22)	刂刀リ刂	虫巴电	罒
口	K(23)	川		
田	L(24)	口	四甲早車酉	皿四串
山	M(25)	冂	由贝几门	冊冊冋冏用网凬
禾	N(31)	丿𠂉	乍𠂉彳攵夂	禾
白	B(32)	丿亻𠂉	气手手斤	
月	V(33)	彡㇏	用乃豸豕豕仄反K	
人	C(34)	八	亻	亼亼𠂎𠂎
金	X(35)	𠂊勹	鱼乌夕	钅犭乂钅夕
言	Y(41)	丶	文方主	亠亩方玄讠
立	U(42)	冫丷⺌	六广辛	亠
水	I(43)	氵	不古	氺水氺水
火	O(44)	灬	小米业	小⺌业小小
之	P(45)	礻	又力冖宀	辶廴マス
心	T(51)	乙	刁雨儿兀	忄⺗丬爿片
子	R(52)	凵凶	孑卩尸耳巳也	孑
女	E(53)	巛儿	刀九臼艮彐	艮
巳	W(54)	コム	己巳尸马厶	尸[illegible]艮艮
乡	Q(55)	五𠂊	弓匕	匕纟幺

Fig. 4

GB 2 158 626 A

name of strokes	form of strokes	code names of strokes	key-code	direction of strokes	similar strokes
horizontal	—	1	G(11)	→	— ㇐ ㇀
vertical	丨	2	H(21)	↓	丨 丨 亅
left-falling	丿	3	N(31)	↙	丿 丿
right-falling	丶	4	Y(41)	↘	丶 丶
turning	乙	5	T(51)	Z	乙 𠃌 𠃍 𠃋

Fig. 1

2158626

2158626

pattern of characters	code name of patterns	topological patterns and examples
left-right	1	江 湘 语 部
up-down	2	节 意 花 华
embracing	3	因 同 床 这 司 区
single	4	重 本

Fig. 2

2158626

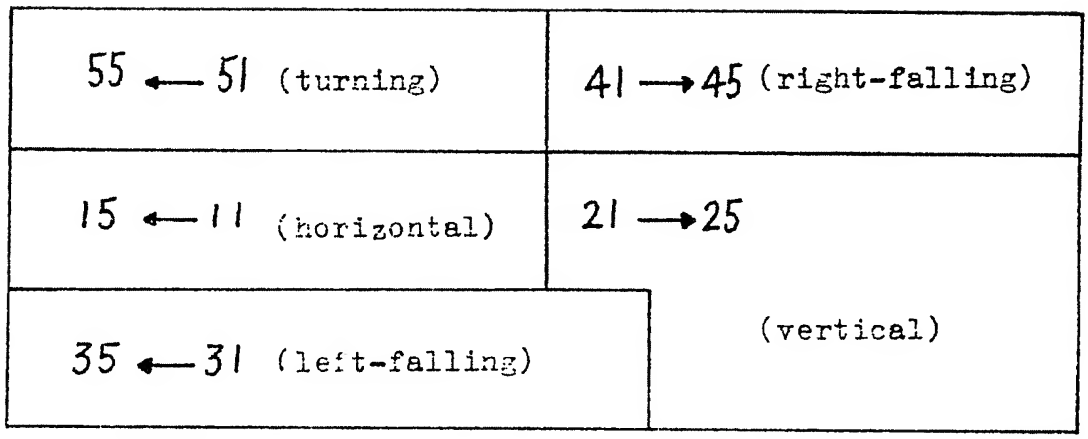


Fig.3

2158626

key names	key-code	stroke-root	main roots	derivative roots
王	G(11)	一	干千夫戋圭壬	垂
土	F(12)	二	士十丰毛	
大	D(13)	三丆ナ	厂石戊尢甘甚井	犬戶非⺕丰手镸
木	S(14)		寸西	柬
廾	A(15)	七匚丂	丁工开戈	廿廿云弋
目	H(21)	丨卜⺊	上止卢	广⺊且
日	J(22)	刂丌刂刂	虫巴央	四
口	K(23)	川		
田	L(24)	囗	四甲里車酉	皿四串
山	M(25)	冂	由贝几门	皿冊冋冏冋冋凸
禾	N(31)	丿𠂉	乍𠂉彳攵夂	禾
白	B(32)	𠂊亻⺈	气手手斤	
月	V(33)	彡⺈	用乃豸豕豕KKK	
人	C(34)	八	亻	𠆢亼癶癶
金	X(35)	勹勹	鱼乌夕	亻犭乂钅夕
言	Y(41)	丶	文方主	亠古亥去讠
立	U(42)	冫丷䒑	六广辛	亠
水	I(43)	氵	不古	氵氺以永
火	O(44)	灬	小米业	⺌丷业小⺗
之	P(45)	礻	又力冖宀	辶廴マス
心	T(51)	乙	习雨儿兀	忄⺗丬爿片
子	R(52)	凵凶	孑卩卩耳已也	孓
女	E(53)	巛儿	刀九臼艮彐	艮
巳	W(54)	コ厶	己巳尸马厶	尸丗吕阝
纟	Q(55)	五夕	弓匕	匕纟幺

Fig. 4

block-number / section-number	1	2	3	4	5
1	王 一	土 二	大 三	木	艹
2	目 丨	日 川	口 川	田 川	山
3	禾 丿	白 ノ	月 川	人	金
4	言 丶	立 丷	水(氵) 氵	火 灬	之
5	心 乙	子 巛	女 巛	已	纟

Fig. 5

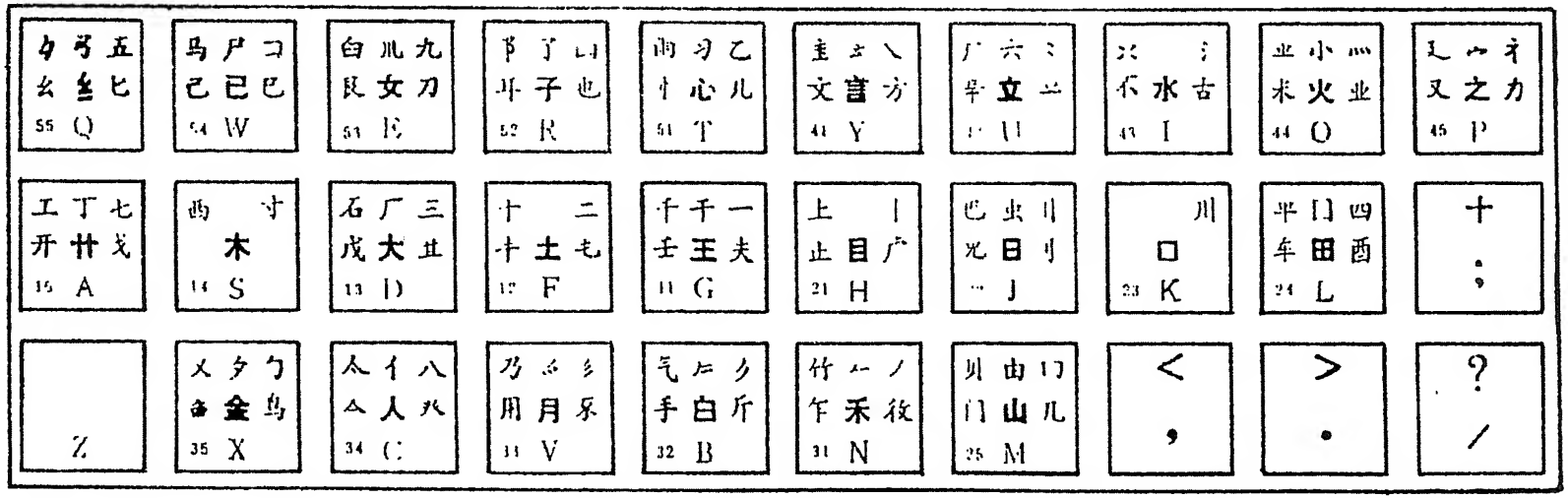


Fig.6

2158626

2158626

the last stroke & its code name \ pattern & its code name	1	2	3	4
— 1	11 (G)	12 (F)	13 (D)	14 (S)
| 2	21 (H)	22 (J)	23 (K)	24 (L)
/ 3	31 (N)	32 (B)	33 (V)	34 (C)
\ 4	41 (Y)	42 (U)	43 (I)	44 (O)
乙 5	51 (T)	52 (R)	53 (E)	54 (W)

Fig. 7

2153626

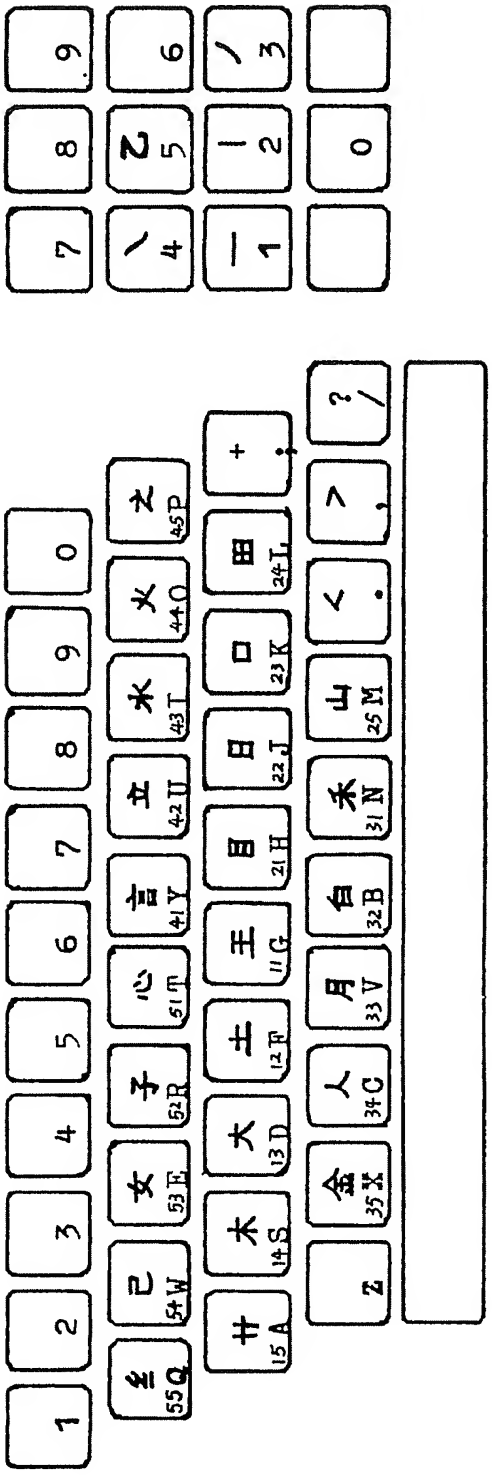


Fig. 8

SPECIFICATION

Encoding Chinese characters and keyboard therefor

Background to the Invention

This invention relates to a universal system for encoding Chinese (or the like) characters and a kind of keyboard designed on the basis of the system. The system is described herein in relation to Chinese characters, but it is applicable to like characters such as Japanese characters.

In a Chinese information processing system, the quick input of individual characters and Chinese words and phrases is a major problem that calls for urgent solution in the extended use of computers in the countries and areas using Chinese characters. Prior art systems for encoding Chinese characters include such methods as encoding by stroke-form, encoding by whole characters at a large keyboard and encoding by pnonemes and morphemes combined. Some such systems encoded Chinese characters only on the basis of single stroke forms; others used their combinations or roots and often needed large numbers of roots without considering their inner links. The systems have disadvantages such as large numbers of keys, the need for special input devices, many rules to be remembered by the operator, complicated operations, too many identical codes and low input speeds. In general, any user of the systems needs a special, long training in order to memorise their rules.

The Invention

The present invention, for the purpose of creating a direct and easy-to-learn method, adopts a system of breaking Chinese (or the like) characters down into roots and encoding them by spelling out the component roots. Using the present invention, Chinese characters, words and phrases can be quickly input entirely according to the information of the character form, and the input is therefore suitable for the entry into any large, medium, small and mini sized computers as well as for Chinese information processing and communications systems. Hence the method of encoding Chinese characters and phrases is compatible with high input speed systems. The roots optimally selected are arranged on keys of a standard keyboard for Chinese; several tens of thousands of characters and words can be entered into and stored in a computer; the operator can key in texts at a speed of about 120-150 characters per minute without looking at the keyboard and the efficiency of inputting characters can be raised enormously.

Using the invention, Chinese characters are regarded as a form of spelling made up of geometrical elements, and can be broken down at three levels, namely, those of stroke forms, roots and whole characters. Strokes are combined to form roots and roots to form characters. The system does not consider the pronunciation, nor does it break every character down into single strokes, but forms characters and words by roots or their codes which can be input in the order of writing the character or word by hand.

For the realisation of the system of the invention, tens of thousands of Chinese characters have been broken down into all their roots and statistics made; the roots have been sorted according to the frequency distribution of the roots in characters and the frequency distribution of the roots in running texts taken from different kinds of newspapers, books and magazines. After the computation of large numbers of roots, the optimally selected roots were sorted and their frequency distributions form an important foundation to the present encoding system.

In an alternative system, which can be used in association with the above system, a keyboard can be used having keys corresponding to the strokes constituting Chinese characters, a predetermined number of the keys being operated in a sequence corresponding to the sequence of writing the character by hand.

The Drawings

The invention is further described, by way of example, with reference to the accompanying drawings, in which:

Figure 1 shows the numerical codes of the five kinds of strokes of Chinese characters;

Figure 2 shows the four topological patterns of Chinese characters and gives examples;

Figure 3 is a chart of blocks in which the roots are sorted in accordance with the numerical codes of the first stroke and the stroke number as well as the frequency with which they are used;

Figure 4 is a chart of all the selected roots;

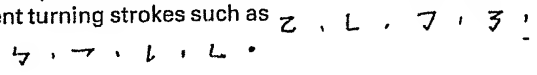
Figure 5 is a chart in wnich the numerical codes of the sections and blocks are shown;

Figure 6 is a diagram of a keyboard for Chinese in which the roots and their numerical codes are compatible;

Figure 7 shows a cross-code of the last stroke and the word's pattern with indentification ability; and

Figure 8 is a schematic drawing of a keyboard for Chinese in which five numerical keys are designed to represent five strokes.

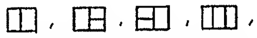
Stroke Code

The basic strokes constituting the roots can be generalised into Horizontal, Vertical, Left-falling, Right-falling and Turning. Each kind of strokes includes those similar in form according to the direction in which they are written. The five kinds of strokes are then allotted the Arabic numerals 1, 2, 3, 4, 5 as their respective codes (as shown in Figure 1), in accordance with their frequency of use. For example the "Turning" kind includes several different turning strokes such as 

Pattern Code

In addition, character forms can be classified into four topological patterns, namely Left-right, Up-down, Embracing, Single. These four patterns are distinguished according to their visual effects and the positions of the roots in characters, and are alloted the Arabic numerals 1, 2, 3, 4 as their respective codes (as shown in Figure 2) in accord-

ance with their frequency. For example, the topological patterns of 江，悟，部，湘，

are 

all called "Left-right" pattern, and are allotted the code 1.

Block Code

The frequency of use of roots in constituting characters was determined by counting the number of times the various roots appeared in a dictionary of all commonly used characters (about 11000). The frequency of use (practical usage) of the roots was determined by counting the number of roots in samples of running texts taken from publications. The weight average value of the two said frequencies determined whether the root was chosen for the keyboard or omitted.

The roots selected should form as many Chinese characters as possible; also, the number of roots chosen must not exceed the capacity of the twenty-five keys of the keyboard. In the specific selection exemplified here, in an initial stage, 800 roots were chosen from the roughly 11000 characters and through the calculation of the frequency distribution of these roots in constituting characters, 250 roots were further chosen from the 800 roots since each of the 250 roots appears at least 10 times in the roughly 11000 characters. It was assumed that the 250 roots are the main roots which may form nearly all characters. However as each key may key in approximately six roots, twenty five keys have in total 150 roots, which are less than '250'. Therefore, at the second state, a calculation was made of the frequency distribution of the 250 roots in practical usage and some of the roots were repeatedly rearranged by incorporating or separating the roots. The remaining 70 roots were classified into derivative or associated roots, as shown in Figure 4.

According to the frequency of use in constituting characters and the frequency of use, the optimally selected roots are then put into five sections which are allotted the numerical codes 1, 2, 3, 4, 5, in accordance with the compatible relations (see below) between them, and each of the five sections is further divided into five blocks which are allotted the Arabic numerals 1, 2, 3, 4, 5, as their block codes (see Figure 5). Thus, there are altogether twenty-five blocks with each having a numerical code of two digits that repesenting its location. Of the two-digit code the first digit, i.e. the section code, it is correspondence with the code of the first stroke of the roots in the section (see Figure 3), and the second digit, i.e. the block code, is either the code of the second stroke of the root in the block, or is the number of strokes in the root. Figure 4 is a chart or diagram and indicates the inner limbs and compatible relations between the roots. The roots associated with any one key have a common characteristic when written, or are of the same origin, or are similar in pattern, in other words they have an "inner link" or "internal connection". The roots on the various keys are compatible in the sense that although any one key represents a number of different roots, the whole code of a character (which is normally entered with four key strokes) is in most cases unique to that character. For example, on the letter key G, they are five roots and a stroke, '干', '千', '壬', '王', '夫',

and '一'. The roots '干' and '千' have the same origion, and so do the roots '王' and '壬'. According to the order of handwriting, these roots have the same or similar direction when the first stroke or the second or the third is written. These 'internal relations' among roots are taken as a criterion in sorting the selected roots into the five blocks as well as into the twenty five groups in order to form the keyboard. In other words, if four letters, for example, E, K, J, O are randomly taken and strand in a certain order as an alpha code, for example, the order of JEK0, where each letter corresponds to a key or a group or different roots (see Figure 4 or 6) as follows:

the letter J corresponding to '曰', '虫',
'刂', '早', '日', 'リ',

the letter E corresponding to '月', '氺',
'九', '艮', '女', '刀',

the letter K corresponding to '口', '川',

the letter O corresponding to '业', '小',
'灬', '米', '火', '业',

the above alpha code JEKO can produce one and only one character, the character '照'; the other possible combinations or permutations of these roots in this alphabetic order would result in nothing but an empty code.

The main roots in the chart of Figure 4 are then arranged onto twenty five keys (from letter A to letter Y) so as to form a chart with twenty five groups of roots. In each group, the typical root with the highest frequence of use is chosen to represent the name of the key (as shown in Figure 4 or 6).

The above selected twenty five groups of roots are arranged on twenty five keys of a standard keyboard in accordance with their probabilities of usage and a habitual or standard fingering in order to form a keyboard for Chinese, as shown in Figure 6. The English letters have no specific function other than assisting those who are familiar with the standard English keyboard and who can operate the keyboard by remembering the English letter equivalents to Chinese characters or roots. Each key has a unique binary code (A.S.C.I.I.).

Encoding the Characters

According to the invention, one can encode or key in a character in units of root in the order of handwriting, and each character has at most four codes. In the simpler form, each code is a single alpha key stroke (which will feed into the computer a single binary coding). For a character which has four roots, all the roots can be keyed in in the order in which they would be written by hand; and for a character whose roots exceed four, one need only

key in the first three roots and the last root. When the code has been input, the computer searches for the character in a whole-code base. When the character is found, it is displayed on the screen current line, and also the character is displayed in the screen prompt line together with its numerical and letter codes.

For example:

The character (照) can be broken down into 日 , 刀 , 口 , 灬 .

The character (赣) can be broken down into the required roots 立 , 日 , 十 , 贝 .

For a character with less than four roots, a "cross-code of last stroke and word's pattern" with identification ability should be added after the code of the roots in the process of encoding and keying in. This cross-code is the combination of the code name of the last stroke and that of the pattern of the character, as shown in Figure 7. The same numerical code corresponding to the cross-code can be found on the keyboard. For example:

The character (汉) can be broken down into 氵 , 又 . Its code is (43, 45), plus its cross-code (41), in which the first digit of (41) refers to the code name of the last stroke and the digit 1 to that of the pattern of the character.

The character (字) can be broken down into 宀 , 子 . Its code is (45. 52), plus its cross-code (12).

The character (本) can be broken down into 木 , 一 . Its code is (14. 11) plus its cross-code (14).

For a character used as the name of a key, the code is to operate its key four times in succession. For example:

The character (土) is used as the name of the key F (12), so its code is FFFF (12. 12. 12. 12.).

The character (言) is used as the name of the key Y (41), so its code is YYYY (41. 41. 41. 41.).

For those roots on the keys (except ones which are the names of the keys) which are in themselves Chinese characters, the process of encoding or keying in is (the code of its key + (the code of its first stroke) + (the code of its second stroke) + (the code of its last stroke). For example:

The character (方) is on the key of letter Y (41), and so the first stroke is (丶) (Y, 41), the second stroke (一) (G, 11), the last (刀) (T, 51); and therefore its input code is YYGT (41. 41. 11. 51.).

As explained above, all single-patterned characters except those on the keyboard are broken down into the roots available on the keyboard and then fed in. When a single-pattern character becomes one part of another character, at most the first two roots of the single-pattern character are included in the encoding. For example:

The character (丙) can be broken down into three roots, i.e. 一 , 冂 , 人 . Its code is GMCO (11. 25. 34. 44).

The character (夷) can be broken down into three roots, i.e. 一 , 弓 , 人 . Its code is GQCO (11. 55. 34. 44).

The character (柄) can be broken down into four roots, i.e. 木 , 一 , 冂 , 人 . According to the above principle, only 木 , 一 , 冂 , and the cross-code of the last stroke and the character's pattern 41 are adopted. Its code is SGMY (14, 11. 25. 41).

Abridged Codes

650 characters have abridged codes, which means one can feed in only the codes of the first two roots of such a character plus an operation on the space key. There are several thousand more characters which can be fed in by feeding in the codes of the first three roots plus an operation of the space key. If an abridged code is input, the computer will search for the character in an abridged-code base, which takes a shorter time than searching for a character in the whole-code base and makes the system more efficient. However, should operator forget the abridged code (or if there is no abridged code), he may input the whole code of a character.

Encoding Phrases

To solve the problem of low input speed, this system adopts a method of encoding Chinese words and phrases that will be explained hereinafter. Therefore thousands of Chinese words and phrases composed by two or more characters are stored to meet the needs of different fields. The two types of codes, the code of a character and that of Chinese words and phrases are compatible in the system. In the case of inputting any phrase, four operations of the key are needed. In the process of inputting characters and phrases alternately, there is no need to shift or add a special operation. To encode a phrase comprising two characters, one can input only the codes of the first two roots of each character. Take the phrase (数学) for example, only the four roots 米 , 女 , 氵 , 冖 are needed and the code of the phrase is OEOP (44. 53. 44. 45); for a phase of three characters, the encoding comprises the code of the first root of each of the first two characters and the codes of the first two roots of the third character. For the phrase (操作员) for example, only the roots 扌 , 亻 , 口 , 贝 are needed and the code of the phrase is BCKM (32. 34. 23. 25.). For a phrase of four characters, only the code of the first root of each of the four characters are needed. For example: for the phrase 汉字编码 , the four roots 氵 , 宀 , 纟 , 石 , are

needed, and the code of the phrase is IPQD (43. 45. 55. 13.). For a phrase of five characters or more, its encoding consists of the codes of the first roots of the first three and the last characters. Take the phrase 中华人民共和国 for example; only the four roots 口, 亻, 人, 口, are needed and its code is KCCL (23. 34. 34. 24.).

Identical Codes

There are very few identical codes in this system of encoding characters. In case of identical codes, the more frequently used character will first appear at the right-hand position of the screen current line. If this character happens to be what the operator needs, he can just go on to the following text and the character would automatically remain in position. If not, he need only operate the space key to exchange that character for another one, which will be shown in the screen prompt line together with its numerical and alpha codes.

Alarm Signals

There are two kinds of alarm signals, a long pip and a short pip. The short one shows that there is no character corresponding to the input code, i.e. an empty code; at the same time, the cursor stops moving. The long one shows that there is an identical code and the operator can handle it with the method just described.

Reserved Key

When keying in the whole of a character, the operator can key in those roots he knows and use the reserved key Z instead of inputting a root of which he is unsure; he can then choose the right character from those shown in the screen prompt line, the prompt line also indicating the full code of the character. Thus as long as one root is keyed in, the character will be shown. In detail, the first five possible characters are shown. If the correct character is shown, it is selected by depressing the appropriate numeral key (one to five) situated above the letter keys on a standard keyboard. If the correct character is not shown, operation of the spacer key brings up the next five characters.

Using Numerical Keys

One can alternatively input any character or Chinese word or phrase by using the five numerical keys 1, 2, 3, 4, 5 on the keyboard, that is to input the numerical codes corresponding to the roots of a character or a phrase.

Order of Inputting Roots

In this system, the operator does not need to care about the order of inputting the roots of a character if he is uncertain about which root should be fed in first, and the result is the same. For example, the character (桃) can be keyed in the order of 木, 𠂉, 儿, , or that of 木, 儿, 𠂉, , and both ways give the same character.

Encoding characters using five strokes

In order to make a beginner master the locations of the roots on a keyboard for Chinese and overcome possible problems when breaking down a character into roots, five keys are further designated for inputting characters. The five keys are numerical keys 1, 2, 3, 4, 5, respectively representing the above mentioned five strokes of Horizontal, Vertical, Left-falling, Right-falling and Turning (一, 丨, 丿, 丶 and 乙) situated on the right-hand position of a standard keyboard, as indicated in Figure 8.

When these keys are used to input a character, the first, the second, the third, the fourth and the last strokes of the character are keyed in in the sequence of writing the character by hand. For example, the first stroke of the character 其 is 一, the second is 丨, the third 丨, the fourth 一, and the last 丶; accordingly, the numerical keys 1, 2, 2, 1 and 4 are operated in that sequence. Hence the character 其 and other characters having the same numerical code or stroke code, for example the character 共 etc., are displayed with the numerical code and their alpha codes in the screen prompt line; these characters are lined up in accordance with their frequency of use and in count-down sequence. Thus the operator may know from a character's alpha code on which keys its roots are located. As for certain characters, for example 从, whose stroke number is less than five, the numerical key O should be operated after the character has been input. Therefore the numerical code of the character 从 is 34340. In addition, the numerical key 6 can be used instead of keying in a stroke of which the operator is unsure.

When the numerical keys are used to input a character, there may be plurality of characters having the same numerical code displayed or to be displayed in the screen prompt line. After the first five characters have been displayed, other characters, if any, will be brought into the prompt line by operating the space key. When a desired character is displayed in the prompt line, for instance, the third character from the left, it can be fed in in position in the screen current line by operating the numerical key 3 which is above the letter keys on the standard keyboard of Figure 8.

CLAIMS

1. A method of encoding Chinese (and the like) characters, or characters, words and phrases, characterised in that the roots and strokes and patterns of characters are optimally selected in accordance with the frequency distributions of the roots both in constituting characters and in practical usage, and classified according to the characteristics of the strokes and the compatibility of the roots.

2. A method of keying in Chinese (and the like) characters, or characters, words and phrases, comprising operating keys in accordance with a breakdown of the characters into roots, strokes and patterns and a classification of the characteristics of the strokes and the compatibility of the roots.

3. The method of Claim 1 or 2, in which on a keyboard used for keying in, the keys are substantial-

ly indicated as shown in Figure 4 of the accompanying drawings.

4. The method of Claim 1 or 2, in which on a keyboard for keying in, the roots are arranged on twenty five keys of a standard keyboard according to the compatible relations between the roots of characters and with reference to standard keyboard-fingering and the probability of usage in keying in the characters.

5. The method of any of the preceding Claims, in which the characters are broken down into basic strokes generalised into five kinds, i.e. Horizontal, Vertical, Left-falling, Right-falling and Turning.

6. The method of Claim 5, in which the basic strokes are allotted the respective numerical codes 1, 2, 3, 4 and 5 according to their frequencies.

7. The method of Claim 5 or 6, in which each of the five kinds of basic strokes includes a number of different but similar strokes.

8. The method of any of the preceding Claims, in which the topological patterns of characters are generalised into four kinds, i.e., Left-right, Up-down, Embracing, and Singular.

9. The method of Claim 8, in which the patterns are allotted respective numerical codes 1, 2, 3 and 4 according to their frequencies.

10. The method of any of the preceding Claims, in which the code of the last stroke of a character and the code of the pattern of the character are combined to form an additional code.

11. The method of any of the preceding Claims, in which words and phrases and characters are all encoded in the same way according to their forms, their encoding being compatible.

12. The method of any of the preceding Claims, in which not more than four codes are used to encode a Chinese phrase of two or more characters.

13. The method as claimed in any of the preceding Claims, in which the selected roots are classified into five groups according to the characteristics of their first strokes.

14. The method of Claim 13, in which a keyboard used for keying in and the groups are allotted the respective numerical codes 1, 2, 3, 4, and 5 according to their frequencies.

15. The method as claimed in any of the preceding Claims, in which the roots are put in five sections in accordance with the compatible relations between them, and each of the five sections is divided into five blocks according to the combination of the characteristics of the second stroke of the roots and the number of strokes in the root.

16. The method of Claim 15, in which the five blocks are allotted the respective numerical codes 1, 2, 3, 4 and 5 according to their frequencies.

17. The method of any of the preceding Claims, in which on a keyboard used for keying in, characters are used as names of keys, and a character used as the name of a key is input by the operating the key four times in succession.

18. The method of any of the preceding Claims, in which on a keyboard used for keying in, characters are used as names of keys, and the input code of a root which is in itself a character not used as the name of a key is (the numerical code of its key) + (the code of its first stroke) + (the code of its second stroke (+ (the code of its last stroke).

19. The method of any of the preceding Claims, in which on a keyboard used for keying in, characters are used as names of keys, and all single-pattern characters except those used as the name of a key are broken down into the roots available on the keyboard in the order of writing the character by hand.

20. The method of any of the preceding Claims, in which the same code defines more than one character, the arrangement being such that the more frequently used characters appears in position in the current line of a screen which is used, which character will remain if a particular key is not operated before going on to the following text.

21. The method of Claim 20, in which the particular key is the shift key.

22. The method of any of the preceding Claims, in which on a keyboard used for keying in, a particular key can be operated if the operator is unsure of a root, and the correct code is given on a screen which is used so that the correct character referring to the input code can be found in the prompt line.

23. The method of Claim 22, in which the particular key is the Z key.

24. The method of any of the preceding Claims, in which on a keyboard used for keying in, a character or phrase can be keyed in by inputting codes corresponding to the roots.

25. The method of Claim 24, in which numeral keys 1, 2, 3, 4 and 5 are used to input the roots.

26. A method of keying in Chinese (and the like) characters, comprising using a keyboard having keys corresponding to strokes constituting the characters, and operating a predetermined number of the keys in a sequence corresponding to the sequence of writing the character by hand.

27. The method of Claim 26, which the sequence is a predetermined number of strokes in the sequence of writing, followed by the last stroke in the sequence of writing.

28. The method of Claim 26 or 27, in which, if the character is constituted by less than said predetermined number of strokes, a further key is operated after the character has been input, to bring the number of keying operations up to said predetermined number.

29. The method of any of Claims 26 to 28, in which the keyboard has a further key which is operated for a stroke of which the operator is unsure.

30. The method of any of Claims 26 to 29, in which said predetermined number of strokes is five, the strokes being Horizontal, Vertical, Left-falling, Right-falling and Turning.

31. The method of Claim 30, in which seven keys are used to input the characters.

32. The method of Claim 31, in which the seven keys are numerical keys 1, 2, 3, 4, 5, 6 and O, the numerical keys 1, 2, 3, 4, and 5 representing said five strokes.

33. The method of Claim 32, in which the numerical key 6 can be used instead of keying in a stroke of which the operator is unsure.

34. The method of Claim 32 or 33, in which the numerical key O should be operated after keying in a character having less than five strokes.

35. The method of any of Claims 26 to 34, and associated with the method of any of Claims 1 to 25.

36. The method of any of the preceding Claims, in which characters having the same numerical code or stroke code are displayed in the prompt line of a screen with their alpha codes indicating the keys on which roots of the characters located, and an appropriate numerical key is operated to select or feed in a character in the prompt line.

37. The method of any of the preceding Claims, and used to input a computer.

38. The method of any of Claims 1 to 36, and used to input an information processing system.

39. The method of any of Claims 1 to 36, and used to input a communication system.

40. A method of encoding Chinese (and the like) characters, words and phrases, substantially as herein described with reference to, and as shown in, the accompanying drawings.

41. A keyboard for keying in Chinese (and the like) characters, or characters, words and phrases, comprising keys representing a break-down of the characters into roots, strokes and patterns and a classification of the characteristics of the strokes and the compatibility of the roots.

42. The keyboard of Claim 41, and associated with a screen on which the character, word or phrase is represented after inputting.

43. The keyboard of Claim 41, or 42, and arranged to be used in the method of any of Claims 2 to 36.

44. A keyboard substantially as herein described with reference to, and as shown in, Figure 6 or Figure 8 of the accompanying drawings.

45. A computer, information processing system or communication system having the keyboard of any one of Claims 41 to 44.

46. The diagram or chart for a system of encoding Chinese (or the like) characters, or characters, words and phrases, substantially as herein described with reference to, and as shown in, Figure 4 of the accompanying drawings.

Printed in the UK for HMSO, D8818935, 9/85, 7102.
Published by The Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained.